(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

B2

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 March 2001 (08.03.2001)

PCT

(10) International Publication Number
WO 01/16339 A1

(51) International Patent Classification[7]: C12N 15/82, 15/63

(21) International Application Number: PCT/CA00/00977

(22) International Filing Date: 25 August 2000 (25.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/151,147 27 August 1999 (27.08.1999) US

(71) Applicant: UNIVERSITY OF GUELPH [CA/CA]; Reynolds Building, Suite 213, Office of Research, Guelph, Ontario N1G 2W1 (CA).

(72) Inventors: BAILEY, Andrea; 5 Cork Street East, Apartment 2B, Guelph, Ontario, N1H 2W7 (CA). YU, WenjiN; 252, Stone Road West, Guelph, Ontario N1G 2V7 (CA). TUBOLY, Tamas; 186 Edinburgh Road South, Apartment 804, Guelph, Ontario, N1G 2H9 (CA). NAGY, Eva; 7 Brazolot Drive, Guelph, Ontario N1G 4M1 (CA). ERICKSON, Larry; 7 Woodduck Court, Guelph, Ontario N1C 1B1 (CA).

(74) Agents: SECHLEY, Konrad, A. et al.; Gowling Lafleur Henderson LLP, Suite 2600, 160 Elgin Street, Ottawa, Ontario K1P 1C3 (CA).

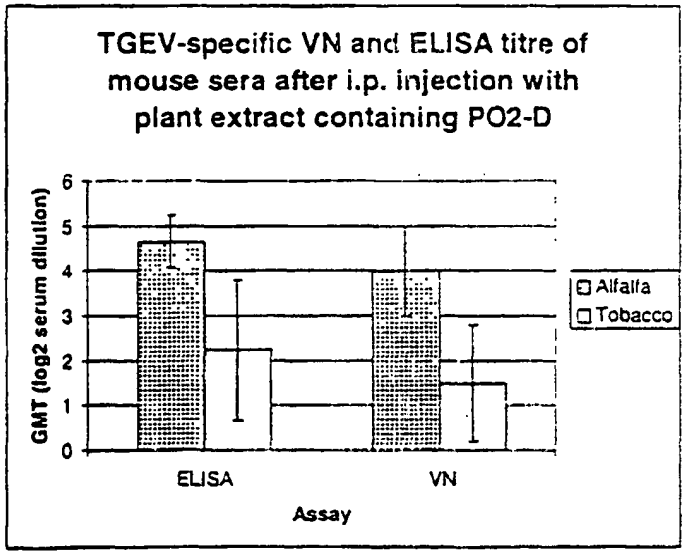
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: USE OF ARABINOGALACTAN PROTEIN FUSION CONSTRUCTS IN A METHOD OF EXPRESSING PROTEINS AND PEPTIDES IN PLANTS



(57) Abstract: The present invention provides a method of expressing a foreign or heterologous protein or peptide in a cell. In particular, the invention concerns a method of increasing foreign protein and peptide expression in a plant cell comprising the use of an arabinogalactan fusion product.

WO 01/16339 A1

USE OF ARABINOGALACTAN PROTEIN FUSION CONSTRUCTS IN A METHOD OF EXPRESSING PROTEINS AND PEPTIDES IN PLANTS

The present invention is in the field of plant molecular engineering. In particular, the invention provides a method to express proteins and peptides in plants using arabinogalactan fusion proteins and peptides.

BACKGROUND OF THE INVENTION

Humans have long looked to plants as a source of medicines, as well as a source of food. Although most medicinal compounds from plants are non-proteinaceous in nature, plants produce a variety of proteins which have immunophysiological functions when ingested (Arai, 1996 Biosci. Biotech Biochem. 60:9). Some examples are the hypotensive effects of a zein peptide from maize and the lowering of blood cholesterol by a glycinin peptide from soybean (Arai, 1996 Biosci. Biotech Biochem. 60:9). The development of transformation techniques has now extended the potential of plants to produce a much broader range of proteins with medicinal/physiological effects, including antibodies, antigens, antibiotics, and growth factors (Owen and Pen, 1996 Transgenic Plants). Since such proteins originate from organisms quite different from plants, they also confer a much wider range of biological activity to plants than is otherwise available.

There are now many examples of "successful" expression of foreign genes in plants, depending on the purpose of expression. Several bacterial enzymes have been expressed at satisfactory levels with very little modification, other than regulatory sequences at the 5' and 3' ends. Well-known examples are the genes for the NPT-II, GUS, CAT, and PAT enzymes. The protein levels needed for an enzyme, however, may be considerably less than what is needed for non-enzymatic proteins. Early attempts to express the Bt gene, from *Bacillus thuringiensis*, in plants are probably more typical of this type of research. Initial constructs containing the naturally-occurring form of the gene, or truncated versions thereof, produced levels of the protein which could not be detected on Western blots, although some biological activity was evident from feeding trials (Vaeck *et al.*1987 Nature 328:33). It was only after extensive re-designing of the gene that adequate levels of the Bt protein were

achieved in plant tissue (Perlak *et al.* 1991 PNAS 88:3324). There are many other examples (and probably many unreported) of expression levels of foreign proteins in plants which are very low, i.e. 0.001% of total protein or less (e.g. Higo *et al.* 1993 Biosci. Biotech. Biochem. 57:1477 and Gomez *et al.* 1998 Virology 249:352).

Several factors leading to low expression levels of foreign proteins in plants have been postulated, including A/T content, codons atypical of plants, premature poly-A signals, mRNA destabilizing sequences and fortuitous intron-like sequences (Koziel *et al.* 1996 Plant Mol. Bio.l 32:393; Gallie, 1996 In: Transgenic Plants (eds. Owen and Pen) pp. 49-74). Some support for these assertions has been provided by increased expression of re-engineered versions of these genes, as in the Bt example. However, there are now probably numerous examples which demonstrate that addressing only these factors may not be sufficient. Experiments with a swine viral antigen and a swine growth factor has shown that the complete re-synthesis of genes with these factors in mind may not result in increases in protein levels in transgenic plants which justify the effort (unpublished results).

Several strategies have been adopted to address the problem of low levels of foreign protein expression in genetically transformed plants. These strategies include targeting the transgene protein to the endoplasmic reticulum (ER), apoplast, or chloroplast, and translational fusions. The role of the signal peptide is now well-established in directing protein synthesis in plants to the ER, as well as the roles of other intrinsic sequences in retaining proteins in the ER (H/KDEL) or targeting them to the vacuole (eg. in various seed storage proteins) (Conceicao and Raikhel 1996 In: Transgenic Plants (eds. Owen and Pen) pp. 75-98). The highest levels of Bt protein (3-5%) reported to date have been synthesized in the chloroplast following integration of the gene in the chloroplast genome (McBride *et al.* 1995. Bio/Technol. 13:362). Wandelt *et al.* (1992. Plant J. 2:181) increased the amount of vicillin by 20 fold in alfalfa leaves by attaching a KDEL sequence to the C terminus of the coding region, thereby retaining the protein in the ER. In contrast, Verwoerd *et al.* (1995. Plant Physiol. 109:1199) achieved the highest levels of phytase in tobacco leaves through secretion to the apoplast.

Expression of foreign proteins and peptides in plants has also been achieved by means of translational fusions to larger "carrier" proteins. A prominent example was expression of enkephalin, a 5-aa neuropeptide, as part of a seed storage protein, reaching levels of 0.1% of total extractable protein (Van dekerckhove *et al.* 1989. Bio/Technol.). Another variation of this approach was the synthesis in plants of a malarial epitope as a part of the coat protein of the tobacco mosaic virus for use as a vaccine (Dalsgaard *et al.* 1997 Nat. Biotech 15:248). Fusions to beta-glucuronidase, a bacterial reporter gene, and green fluorescent protein have been used to not only detect expression of the attached transgene protein, but also determine its location and distribution in tissues (e.g. Parmenter *et al.* 1996 In: Transgenic Plants (eds. Owen and Pen) pp. 261-280). Finally, attachment to an endogenous plant protein was effective for the synthesis in plants of two mammalian proteins with pharmaceutical potential: the human cytokine, granulocyte-macrophage colony-stimulating factor, fused to the signal peptide and first 8 aa of a rice glutelin protein (Sardana *et al.* 1998 In: Recombinant Proteins from Plants (eds. Cunningham and Porter) pp. 77-87); and the human anticoagulant peptide hirudin, produced as an appendage to the 3' end of the oleosin protein (Parmenter et al. 1996 In: Transgenic Plants (eds. Owen and Pen) pp .261-280). In this latter case the fusion to oleosin localized the hirudin to the oil bodies of canola, as the oleosin protein is attached to the membranes of those bodies, which facilitates isolation of the protein by flotation of the oil bodies after grinding of the seed in aqueous buffer.

There is a vast range of mammalian proteins and peptides that one may wish to express in plants for a variety of reasons. Antimicrobial peptides and derivatives thereof could provide plants with some degree of resistance to bacteria and fungi (Hancock and Lehrer, Trends Biotechnol. 16:82;). There is also a major interest in using plants as bioreactors to produce high-value proteins with pharmacological properties (e.g. Parmenter *et al* 1996 In: Transgenic Plants (eds. Owen and Pen) pp. 261-280; Sardana *et al.* 1998 In: Recombinant Proteins from Plants (eds. Cunningham and Porter) pp. 77-87; Cramer *et al.* 1996 In: Transgenic Plants (eds. Owen and Pen) pp.299-310). Another obstacle to the use of plants for such purposes involves not only the production of the proteins, but also the extraction and purification of proteins

and peptides from plant tissues. An option is to orally consume plant tissue in some form to obtain the benefit of the new protein produced therein. Presumably such proteins would be active primarily within the digestive tract (e.g. phytase, antibiotic peptides), affect mucosal tissues and development (e.g. epidermal growth factor), or stimulate an immune response, particularly a mucosal immune response, which would result in the secretion of protective IgA molecules into the intestinal tract.

Despite these advances in plant molecular engineering, there remains many important proteins and peptides that have not proven amenable to high level expression in plants. There is therefore a continuing need for new methods for the expression of foreign proteins and peptides in plants at acceptable levels. There is also a need for a method for the expression of foreign proteins and peptide in plants that allows for the ready extraction and purification of the target protein or peptide.

It is an object of the invention to overcome disadvantages of the prior art.

The above object is met by the combination of features of the main claim, the sub-claims disclose further advantageous embodiments of the invention.

SUMMARY OF THE INVENTION

The present invention provides a method of expressing a foreign or heterologous protein or peptide in a cell. In particular, the invention is directed to a method of increasing foreign protein and peptide expression in a plant cell comprising the use of an arabinogalactan fusion product.

The present invention pertains to a method of expressing a foreign protein or peptide in a cell comprising the steps of:

(a) making a fusion construct comprising one or more regulatory elements in operative association with a chimeric gene, the chimeric gene comprising a first nucleic acid encoding an arabinogalactan protein or a fragment thereof and a second nucleic acid encoding a foreign protein or peptide to be expressed in the cell; and

(b) introducing said fusion construct into a host cell so that the cell expresses the chimeric gene.

Further, the invention provides a method as defined above wherein the first nucleic acid comprises a portion of a gene that encodes an arabinogalactan protein isolated from alfalfa pollen. In particular, the portion of a gene that encodes an arabinogalactan protein is that shown in SEQ ID NO:1 (Figure 1).

In another aspect of the invention, there is provided a method as defined above wherein the host cell is a plant cell.

The present invention also involves recombinant vectors comprising regulatory elements, a portion of a gene that encodes an arabinogalactan protein and a DNA segment encoding the foreign protein or peptide to be expressed in the cell.

The invention also extends to cells comprising any one of the recombinant vectors described above. In particular these cells are plant cells.

The invention also provides for a transgenic plant, or a transgenic seed, the genome of which has been transformed with any of the fusion constructs described above.

This summary of the invention does not necessarily describe all necessary features of the invention but that the invention may also reside in a sub-combination of the described features.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other features of the invention will become more apparent from the following description in which reference is made to the appended drawings in which:

Figure 1 shows the nucleotide and amino acid sequence of alfalfa pollen arabinogalactan (SEQ ID NO's: 1 and 2):

Figure 2 is a graphical representation of the full length alfalfa pollen arabinogalactan protein.

Figure 3 is the amino acid sequence of the fusion protein derived from truncated alfalfa pollen arabinogalactan and the D-epitope of porcine transmissible gastroenteritis virus (TGEV; SEQ ID NO:3).

Figure 4 is a graphical representation of the pBluescript® SK II+ vector.

Figure 5 is the nucleotide sequence of the truncated alfalfa pollen arabinogalactan as subcloned into pBluescript® SK II+ (SEQ ID NO:4). The double underline represents the *HaeII* cleavage site. The recognition sequence for *HaeII* is rgcgc/y, and therefore this restriction enzyme cuts between nucleotides 710 and 711.

Figure 6 shows the oligomers encoding the modified TGEV epitope. Annealing of the two 60-mers (sense, SEQ ID NO:5, and antisense, SEQ ID NO:6) resulted in "sticky ends" at both the 5' and 3' ends.

Figure 7 is a graphical representation of the truncated alfalfa pollen arabinogalactan and the synthetic D-epitope of TGEV.

Figure 8 is a graphical representation of the pBISN1 binary vector used for plant transformation.

Figure 9 is a Western Blot showing expression of the truncated alfalfa pollen arabinogalactan and the D-epitope of TGEV fusion protein in alfalfa (immunobloting with polyclonal antibody from chickens against the truncated alfalfa pollen arabinogalactan). Lane A is the standard low range molecular weight marker; Lanes B-F are protein isolates from 5 transgenic plants; Lane I is an alfalfa tissue culture control.

Figure 10 shows the geometric mean titre (GMT) of TGEV-specific antibodies from serum retrieved from mice immunized with PO2-D in alfalfa (left hand column) or in tobacco (right hand column). Two different assays were done with each serum sample, and a different GMT was calculated for virus neutralization (VN) and for TGEV-recognition (ELISA). The error bars indicate standard deviation.

DETAILED DESCRIPTION OF THE INVENTION

The present invention involves a method of expressing foreign or heterologous proteins and peptides in a cell. In particular, the invention concerns a method of increasing foreign protein and peptide expression in a plant cell comprising the use of arabinogalactan fusion protein and peptides.

The following description is of a preferred embodiment by way of example only and without limitation to the combination of features necessary for carrying the invention into effect.

Arabinogalactan proteins (AGPs) are found in flowering plants from every taxonomic group tested. They constitute a major class of proteins in the extracellular matrix between the pollen and pistil and are also widely distributed in other plant tissues (Chasan, R. Plant Cell 1994, 6:1519; Cheung, A.Y. Proc. Nat. Acad. Sci. USA, 1995, 92:3077; Showalter, A.M. Plant Cell 1993, 5:9). AGPs and their corresponding cDNA's have been isolated from pistils and suspension cells (Chen, C.C. *et al.* Proc. Natl. Acad. Sci. USA 1994, 6:1643; Du, H. *et al.* Plant Cell 1994, 6:1843; Mau, S. *et al.* Plant J. 1995, 8:269). AGPs have also been found in the plasma membrane of pollen vegetative cells (Pennell, R.I. *et al.* J. Cell. Biol. 1989, 108:1967) and sperm cells (Pennell, R.I. *et al.* Plant Cell 1991, 3: 1317), and the cell wall of pollen tubes (Li, Y. *et al.* Planta 1992, 188:532). Two homologous cDNAs encoding putative AGPs were isolated from *Brassica napus* (Gerster, J. *et al.* Plant Physiol. 1996 110:1231).

AGPs that are particularly useful in the method of this invention are those that are, non-membrane bound, non-enzymatic and non-structural. These characteristics make it easier for the AGP fusion protein or peptide to accumulate in the tissue with minimal adverse effects on cell function. However, AGPs that are membrane bound, enzymatic, or structural may also be used. The highly glycosylated nature of AGPs provides these molecules with the tendency to form aggregates. This "sticky" nature

AGPs adds to their potential as fusion proteins as this characteristic may improve isolation and purification of the desired foreign protein or peptide. Such aggregates may also be useful for protecting a foreign protein or peptide from degradation, both in plant tissue and in the digestive tract of animals.

A pollen-derived DNA sequence was cloned as cDNA from an alfalfa pollen-specific library, and sequence analysis of the inferred protein structure indicated it to be an arabinogalactan protein (AGP) (Qiu *et al.* 1997 Plant Sci. 124:41). Interestingly, O-glycosylation is universally found to increase the resistance of the underlying peptide to proteases (Jentoft, 1990 TIBS 15:291), a property that could contribute to the utility of AGPs as fusion proteins and peptides.

The alfalfa pollen AGP is a secreted protein, 259 amino acids long and it has 3 domains (see Figure 2). Domain 1 extends from residue 1 to residue 20, and its sequence matches the consensus sequence of an amino-terminal signal peptide. The presence of such a region is consistent with the observed extracellular location of this AGP (unpublished data). Domain 2 is the central domain of the peptide, and it can be further divided into two subdomains; the first spans amino acids 21-130 and is enriched in aspartate and glutamate, while the second extends from residue 131-238 and is very high in alanine, serine and proline. The O-glycosylation characteristic of AGPs is expected to occur within the region enriched for proline in domain 2 of this AGP. Domain 3 is a hydrophobic region from 239-259, and software analysis of the amino acid sequence suggests that it is a transmembrane helix (Qiu *et al.*, 1997 Plant Sci. 124:41). This third domain has been removed in the fusion construct made here. Many AGPs have, at the carboxy-terminus, a hydrophobic domain. This domain represents an ideal target for removal and replacement with a sequence encoding a foreign protein or peptide of interest to make a chimeric gene to be expressed in a cell or host organism. However, it is to be understood that other fragments of the AGP protein, or other domains of AGP may be used or deleted as required to prepare a fusion protein as described herein. A fragment of AGP comprising at least 6 amino acids of the charged domain (i.e. nucleotides 1-27 of SEQ ID NO2) is desired for the preparation of a fusion protein or peptide as described herein. For example, which are

not to be considered limiting in any manner, a fragment of AGP from about 1 to about 130 (of SEQ ID NO:2; including the signal peptide and charged domain, see Figure 2), or an AGP fragment comprising amino acids 1 to a position within the glycosylated domain (between amino acids 131-238; SEQ ID NO:2; including the signal peptide, charged domain and a portion of the glycosylated domain, see Figure 2) may be used as described herein for the preparation of a fusion protein, however other lengths of the nucleotide seqquence disclosed in SEQ DI NO:2 may be used as needed . A fusion protein or peptide comprising all or a portion of the charged domain (nucleotides 1-130 of SEQ ID NO:2, or a fragment thereof) of AGP may be useful for the expression of cationic peptides, since the charged domain helps enhance the stability of the fusion protein.

All AGPs characterized to date are secreted and therefore ultimately end up in association with the plant cell wall. If expressed within a plant, the secretion of the AGP fusion protein or peptide into the plant cell wall may protect the protein or peptide from rapid degradation in the stomach of animals or humans by trapping the protein or peptide in an "enteric coating" consisting of a matrix of cellulose and hemicellulose microfibrils, lignin and the like. The gradual break-down of the cell wall would result in an extended release of the protein or peptide into the intestinal lumen of the human or animal eating the plant, where it would be needed for some immunophysiological function. However, the AGP fusion construct of the present invention may also be expressed in other host cell expression systems, for example, bacterial, yeast, fungal, insect or mammalian. In some cases, it may be that a small peptide, fused to an AGP or a fragment thereof, must be released from the AGP in order to have biological activity. Such cleavage could be achieved by any means known to one of skill in the art, for example but not limited to proteolytic degradation, for example, the action of trypsin at a trypsin site inserted into the chimeric gene between the AGP and the peptide. The trypsin enzyme is very specific regarding its amino acid cleavage site (following an arginine or lysine) and is abundant in the intestinal tract of humans and livestock animals. It is preferred for a trypsin site to be added only to fusion construct for those proteins or peptides containing no other trypsin sites.

The particulate structure of plant cell wall fragments containing a foreign protein or peptide may also confer an immunological advantage to an antigen produced in plants for the purpose of an oral vaccine, delivered as plant tissue. Such particulate structures are considered important for the generation of a mucosal immune response, important for protection against enteric pathogens. These pathogens also present themselves as particulate structures to intestinal tissues which sample contents of the lumen. Association with cell wall constituents, in a natural or chemically modified state, such as cellulose, may provide the added benefit of adjuvant properties, thereby enhancing the immunogenicity of antigens co-presented in plant tissue as oral vaccines.

The present invention provides a method of expressing a foreign gene in a cell comprising the steps of:

(a) making a fusion construct comprising one or more regulatory elements, a portion of a gene that encodes an arabinogalactan protein (a first nucleic acid) and a DNA segment encoding a foreign protein or peptide to be expressed in the cell (a second nucleic acid); and

(b) introducing said fusion construct into a host cell so that the cell expresses the fusion.

The term "regulatory elements" as used herein is meant to describe DNA sequences which control the initiation, speed and termination of transcription. This may include promoter and termination sequences as are known to one of skill in the art.

The term "promoter" as used herein refers to the nucleotide sequences at the 5' end of a coding region of DNA, or fragment thereof that contain all the signals essential for the initiation of transcription and for the regulation of the rate of transcription. The promoters used to exemplify this invention are constitutive promoters that are known to those skilled in the art. By "constitutive promoter" it is meant a promoter that directs the expression of a gene throughout the various parts of a host, for example a plant and continuously throughout its development. In plants,

examples of such promoters known in the art include, but are not limited to, the synthetic super promoter (Ni, *et al.* Plant J 1995, 7:661, which is incorporated herein by reference) and the CaMV 35S promoter (Odell et al., 1985, *Nature*, 313: 810-812, which is incorporated by reference). The preferred constitutive promoter is the synthetic super promoter. If tissue specific expression of a gene is desired, for example seed or leaf specific expression, then promoters specific to these tissues may also be employed.

Furthermore, as would be known to those skilled in the art, inducible promoters may also be used to regulate the expression of the gene following induction of expression by providing the appropriate stimulus (inducer) for inducing expression. The use of inducible promoters may be required as constitutive synthesis of a foreign protein or peptide may, in some cases, be toxic or inhibit normal growth of the host, for example a plant, with the result that the only plants which regenerate following expression are those expressing low levels of the foreign protein or peptide. In the absence of an inducer, the DNA sequences or genes will not be transcribed. Typically, the protein factor that binds specifically to an inducible promoter to activate transcription is not present or present in an inactive form which is then directly or indirectly converted to the active from by the inducer or inducing treatment. The inducer may be a chemical agent such as a protein, metabolite, growth regulator, herbicide or phenolic compound, a physiological stress imposed directly by heat, cold, salt or toxic elements or indirectly through the action of a pathogen or disease agent such as a virus. A plant cell containing an inducible promoter may be exposed to an inducer by externally applying the inducer to the cell or plant such as by spraying, watering, heating or similar methods.

The chimeric gene construct of the present invention can further comprise a 3' untranslated region. A 3' untranslated region refers to that portion of a gene comprising a DNA segment that contains a polyadenylation signal and any other regulatory signals capable of effecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by effecting the addition of polyadenylic acid tracks to the 3' end of the mRNA precursor. Polyadenylation

signals are commonly recognized by the presence of homology to the canonical form 5' AATAAA-3' although variations are not uncommon.

Examples of suitable 3' regions for use in plant cells, are the 3' transcribed non-translated regions containing a polyadenylation signal of *Agrobacterium* tumor inducing (Ti) plasmid genes, such as the nopaline synthase (*Nos* gene) and plant genes such as the soybean storage protein genes and the small subunit of the ribulose-1, 5-bisphosphate carboxylase (ssRUBISCO) gene. The 3' untranslated region from the structural gene of the present construct can therefore be used to construct chimeric genes for expression in plants.

The chimeric gene construct of the present invention can also include further enhancers, either translation or transcription enhancers, as may be required. These enhancer regions are well known to persons skilled in the art, and can include the ATG initiation codon and adjacent sequences. The initiation codon must be in phase with the reading frame of the coding sequence to ensure translation of the entire sequence. The translation control signals and initiation codons can be from a variety of origins, both natural and synthetic. Translational initiation regions may be provided from the source of the transcriptional initiation region, or from the structural gene. The sequence can also be derived from the regulatory element selected to express the gene, and can be specifically modified so as to increase translation of the mRNA.

To aid in identification of transformed host cells, the constructs of this invention may be further manipulated to include selectable markers. Useful selectable markers include enzymes which provide for resistance to an antibiotic such as gentamycin, hygromycin, kanamycin and the like, or resistance to herbicides such as glufosinate. Similarly, enzymes providing for production of a compound identifiable by colour change, such as GUS (β-glucuronidase), or luminescence, such as luciferase, are useful.

The portion of a gene encoding an arabinogalactan protein (first nucleic acid) is meant to encompass any gene encoding an arabinogalactan protein, of a fragment thereof, that is capable of acting as a fusion protein for the expression of a foreign protein or peptide in a cell, in particular, a plant cell. Preferably, but not to be considered limiting in any manner, the arabinogalactan protein should have a non-enzymatic or non-structural function and be non-membrane bound. It is also desirable for the protein to have a hydrophobic region at the carboxy-terminus. The carboxy terminus may be removed and substituted with the foreign protein or peptide (the second nucleic acid) to be expressed in the cell using conventional molecular biology techniques known to those skilled in the art. It is to be understood, however, that other portions of the AGP may be removed to produce an AGP fusion protein. More preferably, which should not be considered limiting, the AGP is a secreted protein found in the pollen. Most preferably, the portion of a gene encoding an AGP is an alfalfa pollen AGP, in particular, comprising the nucleotide sequence shown in SEQ ID NO:1, and the amino acid sequence shown in SEQ ID NO:2 (Figure 1). However, it is to be understood that fragments of the nucleotide and amino acid sequences shown in SEQ ID NO's 1 and 2 may be used as a first nucleic acid. For example, which is not to be considered limiting in any manner, a fragment of AGP comprising amino acids 1 to about amino acid 130, or amino acid 1 to about amino acid 237 (of SEQ ID NO:2) may be used as described herein, however, other fragments of the AGP protein may also be used.

The DNA segment encoding a foreign protein or peptide to be expressed in the cell may be any gene or DNA segment encoding any protein or peptide desired to be produced in the cell, in particular a plant cell. The term "foreign protein or peptide", also referred to as a "second nucleic acid", is meant to encompass any protein or peptide that is not normally expressed in the target cell. Examples of proteins or peptides that might be the target of expression in plant cells, include, but are not limited to, proteins or peptides of pharmaceutical interest, such as antibodies, antigens, antibiotics, growth factors, hormones, lymphokines and activators. Also included are protein or peptides which are produced in a plant to enhance an agronomic or industrial trait, for example, but not limited to, herbicide resistance and

stress tolerance. Particular proteins or peptides that can be expressed in a plant cell using an arabinogalactan fusion protein include animal viral antigens, for example, but not limited to, an antigenic portion of porcine transmissible gastroenteritis virus (TGEV). This animal viral antigen may be expressed in a plant cell as an AGP fusion product, for example, as a modified D-epitope of TGEV, as shown in Figure 6 (SEQ ID NO;s 5 and 6).

The present invention therefore provides a DNA comprising one or more regulatory elements in operative association with a chimeric gene, the chimeric gene comprising a first nucleic acid encoding an arabinogalactan protein or a fragment thereof, fused with a second nucleic acid encoding a foreign protein or peptide. Furthermore, this invention includes a vector comprising the DNA just described.

The use of the fusion system of the present invention might be used to enhance the levels of foreign proteins or peptides produced at low levels in a host, for example a plant. Within plants, the 5' end of the AGP gene contains naturally-occurring signals for the initiation of transcription and translation, which are major limiting steps in foreign gene expression in plants. However, it is to be understood that other regulatory regions, including promoters and enhancers may be used as desired with the AGP gene of the present invention in order to optimize expression of the AGP fusion protein. Of much greater significance, however, is the probable necessity of such a system for the synthesis of small peptides which may range in size from, say 6 amino acids to perhaps 30 amino acids. It is very unlikely that peptides such as these could be produced in several hosts, for example plants, or that they would be stable, unless fused to a larger, stable protein. Such peptides, when produced in their natural hosts, are synthesized as part of a larger precursor protein, which is subsequently cleaved to yield the much smaller, biologically-active peptide. An example is glucagon-like peptide-2 (GLP-2), a 33-amino acid peptide, produced in humans as part of a much larger protein from which it is cleaved in specific tissues at certain times. The nucleic and amino acid sequences for GLP-2 are known (Drucker, D. Diabetes 1998, 47:159-169). It is highly unlikely that satisfactory levels of such peptides could be produced in plants without a fusion system such that proposed here.

It should be noted that GLP-2 does not contain a trypsin site within its sequence, making it a good target for the incorporation of such a site between the AGP gene and the peptide gene of the fusion construct. Another example of a smaller peptide, ideal for this application is porcine or human epidermal growth factor (EGF) (REF).

It is believed that expression of proteins or peptides using the method of this invention may be produced within a range of host systems, and in the case where a plant is the host this expression is independent of the species of the plant. All plants are believed to express arabinogalactan-type proteins and therefore this strategy should be suitable for all crop species, including monocots and dicots. Particularly, the method is suitable for food and food crop plants. More particularly, the method is suitable for dicots such as alfalfa, clovers, soybean, pea and tobacco.

Plant gene expression vectors can, and have been as described herein, constructed specifically to express arabinogalactan fusion proteins in plant cells. The construction of plant expression vectors must be adapted for use with plant transformation procedures for transferring those vectors into transgenic plants. However, it is to be understood that one of skill in the art may readily adapt the AGP fusion system of the present invention for the expression of fusion proteins of interest within any suitable host system, including bacterial, fungal, yeast, or animal expression systems. By "transformation" it is meant the stable interspecific transfer of genetic information that is manifested phenotypically. For example, which is not to be considered limiting in any manner, the constructs of the present invention can be introduced into plant cells using Ti plasmids, Ri plasmids, plant virus vectors, direct DNA transformation, micro-injection, electroporation, etc. For reviews of such techniques see for example Miki and Iyer, Fundamentals of Gene Transfer in Plants. In *Plant Metabolism*, 2d Ed. DT. Dennis, DH Turpin, DD Lefebrve, DB Layzell (eds), Addison Wesly, Langmans Ltd. London, pp. 561-579 (1997), and more specifically, alfalfa may be transformed according as described in www.oac.uoguelph.ca/www/CRSC/embryo/proced.html. Tobacco leaf discs may be transformed using a protocol described in Fisher *et al.* (1995 Plant Mol. Biol. Rep.

13:278). The present invention further includes a suitable vector comprising the chimeric gene construct.

The chimeric gene product expressed in a desired host may be feed to an animal or extracted and administered to an animal. For example which is not to be considered limiting in any manner, of the fusion product, or the foreign protein or peptide, as described herein, may be introduced into any desired plant, including forage plants, food crops, or other plants depending upon the need. Examples of such plants include, but not limited to, alfalfa, corn, barley, tobacco, and potato. In the experiments outlined below, alfalfa and tobacco have been used as test organisms for the expression of a fusion product, however it is to be understood that the constructs of the present invention may be introduced and expressed in any plant, other host organism, or cell expression system.

It is contemplated that if a plant is the host, then a transgenic plant comprising a heterologous fusion protein as described herein may be administered to an animal in a variety of ways depending upon the need and the situation. For example, if the protein is orally administered, the plant tissue may be harvested and directly feed to the animal, or the harvested tissue may be dried prior to feeding, or the animal may be permitted to graze on the plant with no prior harvest taking place. It is also considered within the scope of this invention for the harvested plant tissues to be provided as a food supplement within animal feed. If the plant tissue is being feed to an animal with little or not further processing it is preferred that the plant tissue being administered is edible. Furthermore, the protein obtained from the transgenic plant may be extracted prior to its use as a food supplement, in either a crude, partially purified, or purified form using method commonly known within the art. In this latter case, the protein may be produced in either edible or non-edible plants.

Expression of a fusion protein of the present invention, for example but not limited to an AGP:D-epitope of TGEV fusion, within a plant host indicates that fusion

proteins produced and expressed as described herein are biologically active and are capable of eliciting an immune response within an animal to which they are administered (see Example 5).

The above description is not intended to limit the claimed invention in any manner, furthermore, the discussed combination of features might not be absolutely necessary for the inventive solution.

The following examples, while illustrating the embodiments of the invention, are not to be considered as limiting the scope of this invention in any manner.

EXPERIMENTAL EXAMPLES

EXAMPLE 1: CONSTRUCTION OF THE GENE ENCODING THE PO2D FUSION PROTEIN

A fusion protein comprising a truncated version of an alfalfa pollen protein, designated as PO2, fused to a modified D-epitope from the S-protein of TGEV was prepared. PO2 was truncated after alanine-237 (see Figure 1), and a 17 amino acid synthetic epitope was added to the carboxy-terminus. The amino acid sequence of the complete fusion protein is shown in Figure 3 (SEQ ID NO:3).

The initial step in creating the gene encoding this fusion protein involved the restriction digestion of PO2 cDNA that had been subcloned into pBluescript SK II + at the *EcoRI* and *XhoI* sites (Stratagene, see Figure 4). PO2 was first excised from the pBluescript vector by restriction digestion with *XbaI* (Life Technologies). The resulting fragment of approximately 820 base pairs was purified from a 1% agarose TAE gel using the Clontech PCRPure kit according to the instructions of the manufacturer. This PO2 fragment was then cut at position 710 by the *HaeII* restriction enzyme (New England Biolabs) (see Figure 5; SEQ ID NO:4), and the desired 800 base pair fragment representing the truncated PO2 was extracted from an agarose gel, once again using the PCRPure kit.

The 5' overhang (sequence GCGC) generated by the cutting of PO2 with *HaeII* was matched by a complementary strand in the DNA encoding the D-epitope. This double stranded DNA fragment was created by annealing two synthesized 60-mers (see Figure 6; SEQ ID NOs: 5 and 6). A mixture of 250 pmols of each oligomer was prepared in a 50 µl reaction volume with 1x PCR buffer (10x buffer for *Taq* DNA polymerase - Life Technologies). This reaction mix was heated at 90-94 °C for 5 minutes and then slowly cooled to room temperature. The final double-stranded sequence contained a 5' *HaeII* 'sticky end' and a 3'*SacI* 'sticky end'. The codons used in this sequence were those found to be preferred in alfalfa (Plant Molecular Biology Labfax, pp.37-40).

The truncated PO2 gene and the annealed oligomers encoding the D-epitope were ligated into the pBluescript vector in an antisense orientation between the *SacI* and *XbaI* sites. To achieve this, pBluescript underwent restriction digestion with *SacI* and *XbaI* (Life Technologies), and was retrieved from a 1% agarose TAE gel using the PCRPure kit. The purified cut vector, along with the PO2 fragment and the synthetic epitope, were ligated with T4 DNA ligase using the procedure for cohesive-end ligation provided by the manufacturer (Life Technologies). Approximately 20 ng of pBluescript, 20 ng of truncated PO2 and 4.5 ng of the synthetic DNA were used in a 20 µl volume, representing a roughly 1:4:10 molar ratio of fragments. The expected ligation product is depicted in Figure 7.

Following the ligation step, 80 µl of calcium chloride buffer (50 mM $CaCl_2$, 10 mM Tris-HCl; pH 8.0) was added to the ligation reaction. This final 100 µl volume was used to transform *Escherichia coli* strain DH5α using the heat-shock tranformation protocol detailed in Current Protocols in Molecular Biology (Ausubel *et al.*, editors). Transformants were selected on the basis of blue-white colour screening on LB plates containing ampicillin (50 µg/ml), X-gal (20 µg/ml) and IPTG (0.1 mM), and screened by colony PCR (procedure derived Lee and Cooper, 1995). The primers used included one specific for the 5' end of PO2 (cgctggccatatgggtttgaaaaacacc; SEQ

ID NO:7) and the reverse primer for pBluescript (Stratagene sequence - ggaaacagctatgaccatg SEQ ID NO:9).

One colony was found to have a PCR product of the expected size, and the plasmid was extracted and analysed by restriction digest. Double digestion with *XbaI* and *SacI* yielded a fragment of the expected length. Sequencing of this using the T3 primer indicated that the 3' end of the construct was as expected. This PO2D gene was cut from pBluescript using *XbaI* and *SacI* and gel purified from a 1% agarose TAE gel (NucleoTrap kit by Clontech according to instructions provided).

EXAMPLE 2: INTRODUCTION OF PO2D INTO PLANT TRANSFORMATION VECTORS

The binary vector pBISN1 (see figure 8 and Ni *et al.* 1995 The Plant Journal, 7:661) was used for plant transformation. In order to introduce PO2D into this vector it was first cut with the restriction enyzmes *SacI* and *XbaI* and then the digestion mixture was run on a 1% agarose TAE gel. The band of approximately 10 kb was excised using the NucleoTrap kit. This fragment was ligated to the PO2D fragment using T4 DNA ligase in a 20 µl volume. *E. coli* DH5α was transformed using 10 µl of this ligation mix and the $CaCl_2$/heat shock procedure described previously. Colony PCR using the 60-mer D-antisense (SEQ ID NO:6; Figure 6) and a primer specific for the superpromoter in pBISN1 (aagaacggaatgcgcgtg; SEQ ID NO:9) indicated that the ligation had been successful.

EXAMPLE 3: TRANSFORMATION OF PLANTS

In order to introduce the PO2 gene into tobacco and alfalfa plants, the binary vector had to first be introduced into an *Agrobacterium tumefaciens* bacteria strain C58CIRif containing the helper plasmid pMP90 (Koncz and Schell, 1986 Mol. Gen. Genet. 204:383). This was achieved using a protocol modified from Holsters *et al.*

(1978 Mol. Gen. Genet. 163:181). The presence of pBISN1 in the *A. tumefaciens* was assayed using colony PCR (D-antisense and superpromoter primers).

One colony containing the modified pBISN1 vector with the PO2D gene was used to transform alfalfa and tobacco plants.

Alfalfa transformation was performed according to a procedure described on the following web page: www.oac.uoguelph.ca/www/CRSC/embryo/proced.html.

Tobacco leaf discs were transformed using a protocol described in Fisher *et al.* 1995 Plant Mol. Biol. Rep. 13:278.

Once transgenic plants were recovered they were potted and introduced into the greenhouse.

EXAMPLE 4: ANALYSIS OF TRANSGENIC PLANTS

The putatively transformed plants were screened using PCR with *Taq* DNA polymerase from Life Technologies (using manufacturer's instructions). Genomic DNA was extracted using a method adapted from Edwards, Johnstone and Thompson (Edwards *et al.*, 1991 Nucleic Acid Res. 19:1349). Positive plants were further analysed for protein expression using Western blots with denatured protein.

Protein from transformed alfalfa plants was extracted by grinding 50-100 mg of plant tissue in 2x volume of TBS buffer (50 mM Tris base, 150 mM NaCl; pH 7.5) in the presence of 1x protease inhibitor (Boehringer Manheim). 6x SDS loading buffer (62.5 mM Tris-Cl, 2% SDS, 5% β-mercaptoethanol, 10% glycerol) was added to a final concentration of 1x. This mixture was incubated at 55-60 °C for 15-20 minutes and vortexed thoroughly before being centrifuged (15000 rpm; 10 minutes; room temperature). The supernatant was transferred to a fresh tube and stored at -20 °C.

Prior to running the protein on an acrylamide gel the solution was boiled at 95 °C for 4 minutes and centrifuged 4 minutes (15000 rpm; room temperature). 10-30 µl of extract was run at 150 V on an SDS-containing 10% polyacrylamide gel made according to a procedure found in Gel Electrophoresis of Proteins (Hames and Rickwood ed. 1981. pp 24-35). After approximately 1 hour, the protein was transferred to a polyvinylidene difluoride (PVDF) membrane. This was done at 4 °C in SDS transfer buffer (25 mM Tris-base, 190 mM glycine, 20% methanol, 0.1% SDS) at 30 V for overnight. A duplicate protein gel was stained with Coomassie brilliant blue and relative amounts of protein per lane estimated.

Blocking of the PVDF membrane was accomplished by incubating the membrane with 3% milk powder in TBS overnight at 4 °C. The primary antibody (polyclonal anti-PO2 IgY from chicken: courtesy of A. Agyare-Tabbi) was prepared in a 1:2000 dilution with 1% milk powder in TBST (TBS with 0.5 µl/l Tween-20), and the blot was incubated in this for 90 minutes at room temperature. This was followed by 3-4 five minute washes in TBST. The secondary antibody used was mouse anti-chicken IgG conjugated to alkaline phosphatase (Sigma, catalogue no. A-9171), this was added as a 1:10000 dilution in 1% milk powder in TBST. Four 10 minute washes in TBST followed and the blot was developed with BCIP and NBT (Alkaline Phosphatase Substrate Package -Life Technologies) according to the instructions of the manufacturer. A representative blot is included as figure 9.

Similar experiments were performed with transgenic tobacco. Preliminary evidence suggests that PO2 and derived fusions are expressed highly in leaf tissue of tobacco.

EXAMPLE 5: ANIMAL STUDIES

Injection and bleeding of mice

Groups of five BALB/c mice were intraperitoneally (i.p.) injected once weekly for four weeks, with 1 mL/mouse of crude soluble protein extract from leaf tissue ground in PBS buffer (see above). Four separate mouse groups were used in this experiment, resulting in a total of 20 mice. Group 1 was injected with PO2-D alfalfa extract, group 2 with PO2-D tobacco, group 3 with control alfalfa, and group 4 with control tobacco. Blood was retrieved from mice using a tail bleeding procedure five days after the last injection. The blood was kept at room temperature for 1 hour until clotted and then centrifuged in a microcentrifuge at maximum speed for five minutes. The serum was kept at -20°C until needed.

Serology

Virus neutralization assay. The virus neutralization (VN) assay was performed using a technique described by Tuboly *et al.* (1994). Serial dilutions of mouse blood sera were made in 100 µL volumes of Dulbecco's Minimum Essential Medium (DMEM) and incubated with 100 p.f.u. of TGE virus (Purdue-115 strain) at 37°C for 1 hour. These dilutions were transferred into a 96-well microtitre plate on which swine testicular (ST) cells had been grown to confluency. The plates were incubated in a humidified environment with 0.5% CO_2, and monitored for the formation of cytopathic effects (CPE). The first signs of CPE appeared between 18 and 24 hours post-inoculation. The neutralizing antibody determined to be the highest dilution in the mixture applied to a well where CPE was not present.

ELISA. The ELISA assay used was described by Cubero *et al.* (1993). It is a fixed cell assay, in which ST cells had again been fixed to a 96-well plate. Here, the cells were infected with Purdue-115 strain of TGEV but washed with PBS and fixed with acetone (80% v/v) before any cytopathic effect had occurred. After washing with PBST (PBS buffer with 0.05% Tween-20) the plate was blocked for 15 minutes at 37°C in 1% milk powder in PBST. The plate was washed three times with PBST, and serial dilutions of antibody in PBST were added to each well, incubated for 1 hour at room temperature, washed three times with PBST, and anti-mouse rabbit IgG-

horseradish peroxidase (HRP) conjugate was added in a dilution required by the manufacturer (Kirkegaard and Perry Laboratories). After incubating the plate for 1 hour at 37°C, the plate was washed with PBST and developed with 2,2'-azino-di-3-ethylbenzthiazoline-6-sulfonate (ABTS). Following a 30 minute incubation in a dark area at room temperature, the reaction was stopped and the OD405 measured. The TGEV-specific antibody titre of the serum was determined to be the highest dilution at which an OD measurement twice as high as the control wells and at least 0.1 was made.

Statistical analysis

Alfalfa and tobacco data were compared by performing a Student's t-test following the determination of a geometric mean titre (GMT) for serum retrieved from injected animals. The geometric titres (GT) used in calculating the mean were the log2 of the inverse of the detected serum titre (e.g. a titre of 1:16 had a GT of 4). Serum from animals injected with extract from control plants did not neutralize virus in the VN assay. A positive response in the ELISA was based on a OD405 twice as high as the negative control and at least 0.1, and no further statistical analysis was performed in the comparison of the pooled control sample to the samples collected from PO2-D vaccinated animals.

Immune response to alfalfa extracts

All the serum extracted from mice injected intraperitoneally with an extract obtained from transgenic alfalfa comprising PO2-D showed detectable antibodies (Table 1). Two of the five mice in the PO2-D group died of shock directly following the injection and no serum was retrieved. Serum from mice injected with control alfalfa extract (group 3) was pooled, and no antibodies capable of neutralizing TGEV were detected. The ELISA OD_{405} values for the control sera were used as a measure of the absence of TGEV-specificity, and all positive dilutions were determined to have an OD twice as high as these numbers and at least 0.1. The virus-neutralizing activity (see above) of TGEV-specific antibodies was determined by an absence of

CPE on the microtitre plate, however after more extended times (2-3 days) cellular infection could be seen.

Table 1: Alfalfa-immunized (Group 1) mice

Mouse	TGEV-specific antibody titre	
	VN	ELISA
1.	1:16	1:20
2.	1:8	1:20
3.	1:32	1:40

Immune response to tobacco extracts

Three of the four serum samples from the mice exposed i.p. to tobacco plant extract showed antibodies detectable in VN and ELISA assays (Table 2). Again, the serum from mice injected with control tobacco (group 4) was pooled before the assay, and no TGEV-specific antibodies were detected. One mouse from the PO2-D injection group died of shock following plant extract administration, and is not represented here. Similarly to the VN assay described in the alfalfa experiment above, after prolonged incubation (>3 days), CPE could be observed.

Table 2: Tobacco-immunized (Group 2) mice

Mouse	TGEV-specific antibody titre	
	VN	ELISA
1.	1:4	1:10
2.	0	0
3.	1:2	1:10
4.	1:8	1:5

A graphical representation of this data is shown below (Figure 10), in which the geometric mean titres of the alfalfa and tobacco serum samples are compared. The difference in TGEV-specific antibody response to the plant extracts is significant, at $P<0.05$ for the VN data and $P<0.02$ for the ELISA data.

While the present application has been described with reference to what are presently considered to be the preferred examples, it is to be understood that the invention is not limited to the disclosed examples. To the contrary, the invention is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims.

All publications, patents and patent applications are herein incorporated by reference in their entirety to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated by reference in its entirety.

WE CLAIM

1. A method of expressing a foreign gene in a cell comprising the steps of:

(a) making a fusion construct comprising one or more regulatory elements in operative association with a chimeric gene, said chimeric gene comprising a first nucleic acid encoding an arabinogalactan protein or a fragment thereof, and a second nucleic acid encoding a foreign protein or peptide to be expressed in the cell; and

(b) introducing said fusion construct into a host cell so that said cell expresses said chimeric gene.

2. The method according to claim 1 wherein said cell is a plant cell.

3. The method according to claim 2, wherein in said step of making, said first nucleic acid encodes an arabinogalactan protein that has a non-structural and non-enzymatic function.

4. The method according to claim 3, wherein said first nucleic acid encodes an arabinogalactan protein that has a hydrophobic tail at the carboxy terminus.

5. The method according to claim 4 wherein said first nucleic acid comprises the sequence shown in SEQ ID NO:1 (Figure 1).

6. The method according to claim 2, wherein said plant is selected from the group consisting of food and food crop plants.

7. The method according to claim 6, wherein said plant is selected from the group consisting of alfalfa, soybean, pea and clovers.

8. The method according to claim 1, wherein in said step of making, said second nucleic acid encodes a protein or peptide of pharmaceutical interest.

9. The method according to claim 8, wherein said protein or peptide is selected from the group consisting of an antibody, antigen, antibiotic, growth factor, hormone, lymphokine and activator.

10. The method according to claim 2, wherein said protein or peptide is selected from the group consisting of a protein or peptide that enhances herbicide resistance of said plant and a protein or peptide that enhances stress resistance of said plant.

11. The method according to claim 9, wherein said protein or peptide is a viral antigen.

12. The method according to claim 11, wherein said viral antigen is a porcine transmissible gastroenteritis virus (TGEV) antigen.

13. The method according to claim 12, wherein said TGEV antigen comprises a modified D-epitope of porcine TGEV encoded by the nucleic acid sequence of SEQ ID NO:4, SEQ ID NO:5, or a combination thereof.

14. The method according to claim 1, wherein in said step of making, said second nucleic acid encodes a peptide comprising about 3 to about 50 amino acids.

15. The method according to claim 14, wherein said peptide is selected from human epidermal growth factor, porcine epidermal growth factor and human glucagon-like peptide.

16. The method according to claim 2, wherein said fusion construct is introduced into said plant cell by *Agrobacterium*-mediated transformation.

17. The method according to claim 1, wherein in said step of making, said fusion construct further comprises a selectable marker.

18. The method according to claim 1 wherein in said step of making, said regulatory regions comprises a promoter effective in a plant cell and a transcriptional terminator effective in a plant cell.

20. The method according to claim 1, wherein in said step of making said fusion construct does not contain a trypsin cleavage site.

21. The method according to claim 1, wherein said fusion construct encodes a trypsin cleavage site located between said first and second nucleic acids.

22. An isolated DNA molecule comprising one or more regulatory elements in operative association with a chimeric gene, said chimeric gene comprising a first nucleic acid encoding an arabinogalactan protein or a fragment thereof, fused with a second nucleic acid encoding a foreign protein or peptide.

23. A vector comprising said isolated DNA molecule of claim 22.

24. The vector according to claim 23 further comprising a DNA sequence encoding a trypsin cleavage site located between said first and second nucleic acids.

25. A cell transformed with said vector according to claim 23.

26. The cell of claim 25, wherein said cell is a plant cell.

27. A transformed plant comprising said vector of claim 23.

28. A transgenic seed comprising said vector of claim 23.

29. A method for the preparation of a peptide or protein for administration to an animal comprising;

i) transforming a plant with a vector as defined in claim 23 to produce a transformed plant;

ii) growing and harvesting said transformed plant; and

iv) obtaining plant tissue from said transformed plants for administration to an animal.

30. The method of claim 29, wherein in the step of obtaining, the peptide or protein of interest is isolated from said transformed plant prior to administration.

31. The method of claim 29, wherein in the step of obtaining, the plant tissue is orally administered to said animal.

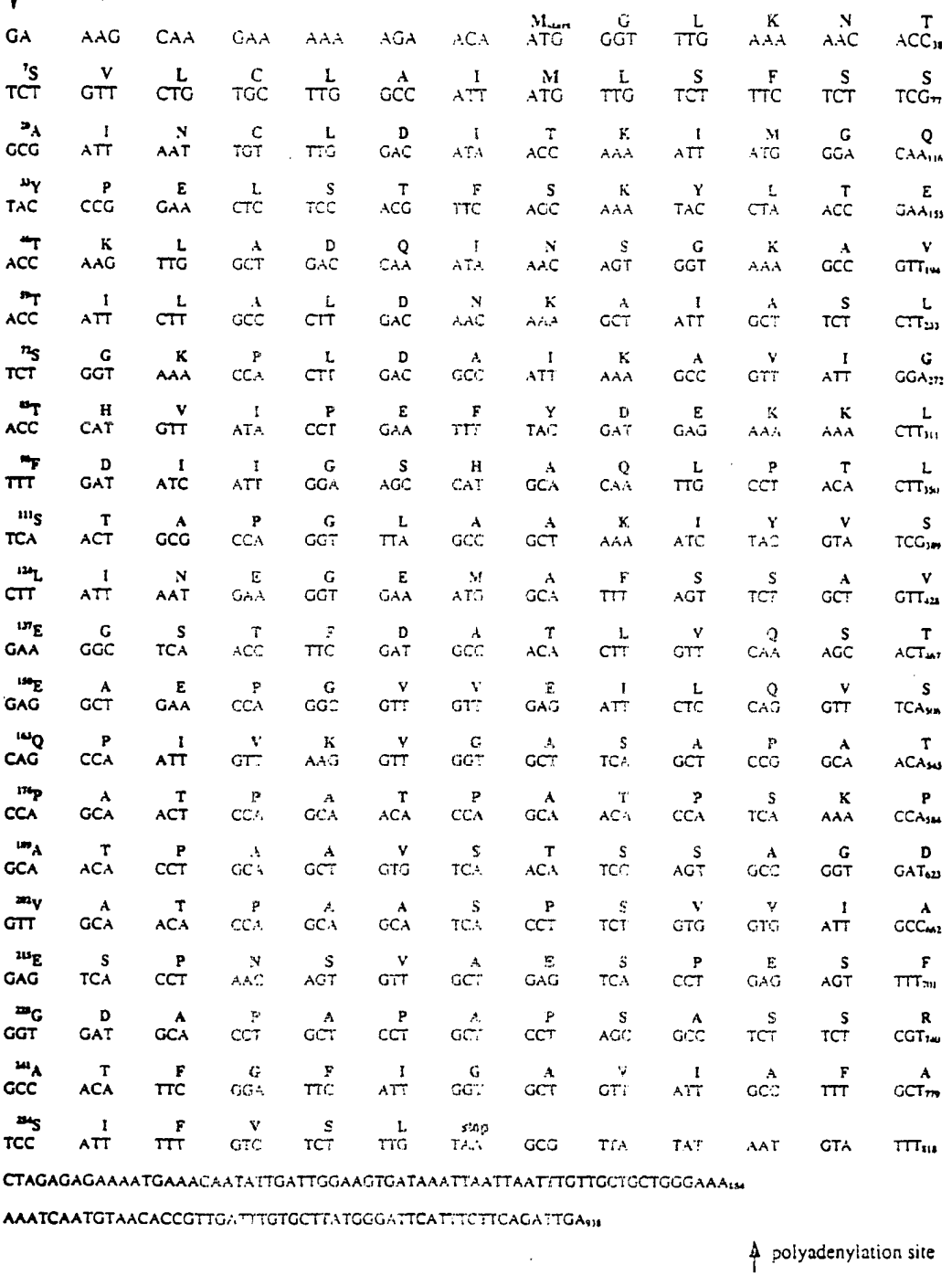

Transcription start site
GA AAG CAA GAA AAA AGA ACA M start ATG G GGT L TTG K AAA N AAC T ACC 38
7 S TCT V GTT L CTG C TGC L TTG A GCC I ATT M ATG L TTG S TCT F TTC S TCT S TCG 77
29 A GCG I ATT N AAT C TGT L TTG D GAC I ATA T ACC K AAA I ATT M ATG G GGA Q CAA 116
50 Y TAC P CCG E GAA L CTC S TCC T ACG F TTC S AGC K AAA Y TAC L CTA T ACC E GAA 155
68 T ACC K AAG L TTG A GCT D GAC Q CAA I ATA N AAC S AGT G GGT K AAA A GCC V GTT 194
89 T ACC I ATT L CTT A GCC L CTT D GAC N AAC K AAA A GCT I ATT A GCT S TCT L CTT 233
71 S TCT G GGT K AAA P CCA L CTT D GAC A GCC I ATT K AAA A GCC V GTT I ATT G GGA 272
85 T ACC H CAT V GTT I ATA P CCT E GAA F TTT Y TAC D GAT E GAG K AAA K AAA L CTT 311
95 F TTT D GAT I ATC I ATT G GGA S AGC H CAT A GCA Q CAA L TTG P CCT T ACA L CTT 350
111 S TCA T ACT A GCG P CCA G GGT L TTA A GCC A GCT K AAA I ATC Y TAC V GTA S TCG 389
124 L CTT I ATT N AAT E GAA G GGT E GAA M ATG A GCA F TTT S AGT S TCT A GCT V GTT 428
137 E GAA G GGC S TCA T ACC F TTC D GAT A GCC T ACA L CTT V GTT Q CAA S AGC T ACT 467
150 E GAG A GCT E GAA P CCA G GGC V GTT V GTT E GAG I ATT L CTC Q CAG V GTT S TCA 506
163 Q CAG P CCA I ATT V GTT K AAG V GTT G GGT A GCT S TCA A GCT P CCG A GCA T ACA 545
176 P CCA A GCA T ACT P CCA A GCA T ACA P CCA A GCA T ACA P CCA S TCA K AAA P CCA 584
189 A GCA T ACA P CCT A GCA A GCT V GTG S TCA T ACA S TCC S AGT A GCC G GGT D GAT 623
202 V GTT A GCA T ACA P CCA A GCA A GCA S TCA P CCT S TCT V GTG V GTG I ATT A GCC 662
215 E GAG S TCA P CCT N AAC S AGT V GTT A GCT E GAG S TCA P CCT E GAG S AGT F TTT 701
228 G GGT D GAT A GCA P CCT A GCT P CCT A GCT P CCT S AGC A GCC S TCT S TCT R CGT 740
241 A GCC T ACA F TTC G GGA F TTC I ATT G GGT A GCT V GTT I ATT A GCC F TTT A GCT 779
254 S TCC I ATT F TTT V GTC S TCT L TTG stop TAA GCG TTA TAT AAT GTA TTT 818
CTAGAGAGAAAATGAAACAATATTGATTGGAAGTGATAAATTAATTAATTTGTTGCTGCTGGGAAA 884
AAATCAATGTAACACCGTTGATTTGTGCTTATGGGATTCATTTCTTCAGATTGA 938
polyadenylation site

FIGURE 1

Alfalfa Potein AGP
NH2 -
- COOH
Signal peptide res. 1-20
Charged domain - asp, glu rich res. 21-130
Glycosylated domain - ala, ser, pro rich res. 131-238
Membrane-spanning domain res. 239-259

Figure 2

```
  1   MGLKNTSVLC  LAIMLSFSSA  INCLDITKIM  GQYPELSTFS  KYLTETKLAD
 51   QINSGKAVTI  LALDNKAIAS  LSGKPLDAIK  AVIGTHVIPE  FYDEKGLFDI
101   IGSHAQLPTL  STAPGLAAKI  YVSLINEGEM  AFSSAVEGST  FDATLVQSTE
151   AEPGVVEILQ  VSQPIVKVGA  SAPATPATPA  TPATPSKPAT  PAAVSTSSAG
201   DVATPAASPS  VVIAESPMSV  AESPESFGDA  PAPAPSASFF  SYGEIQLAKD
251   KVNE
```

Figure 3

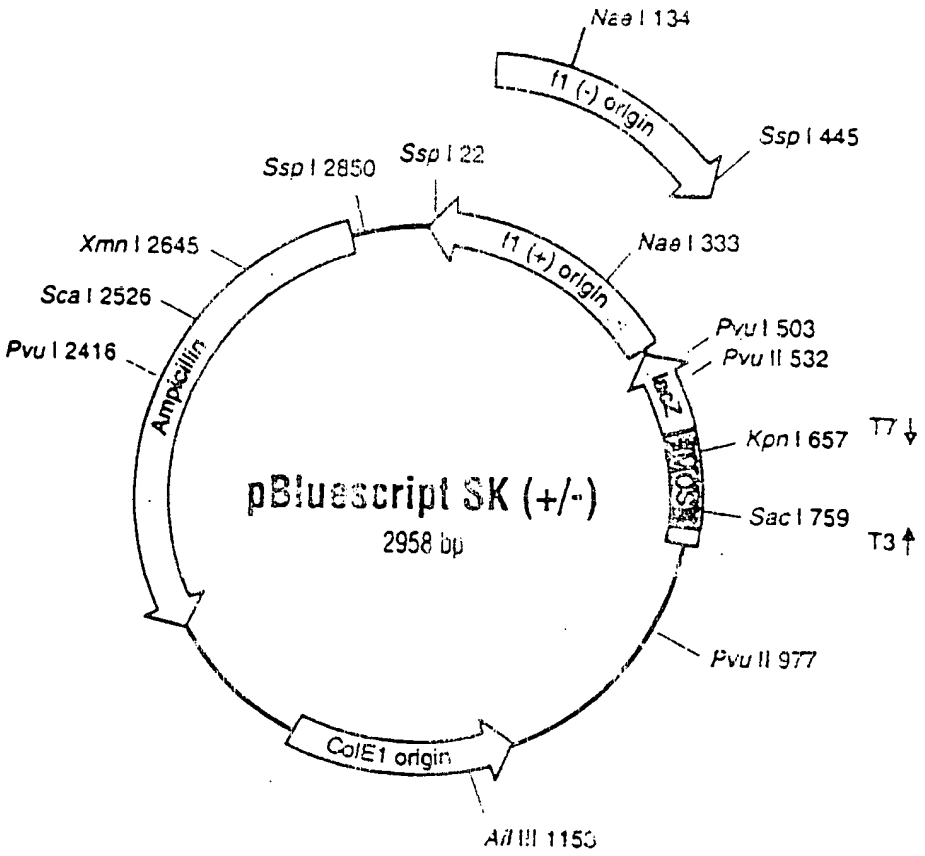

Nae I 134
f1 (-) origin
Ssp I 445
Ssp I 2850
Ssp I 22
Xmn I 2645
f1 (+) origin
Nae I 333
Sca I 2526
Pvu I 503
Pvu I 2416
Pvu II 532
Ampicillin
lacZ
MCS
Kpn I 657
T7
pBluescript SK (+/-)
2958 bp
Sac I 759
T3
Pvu II 977
ColE1 origin
Ava III 1153

Figure 4

```
-32 caatactacg tggaaagcaa gaaaaaagaa caatgggttt gaaaaacacc tctgttctgt
 29 gcttggccat tatgttgtct ttctctttcgg cgattaattg tttggacata accaaaatta
 89 tgggacaata cccggaactc tccacgttca gcaaatacct aaccgaaacc aagttggctg
149 accaaataaa cagtggtaaa gccgttacca ttcttgccct tgacaacaaa gctattgctt
209 ctcttttctgg taaaccactt gacgccatta aagccgttat tggaacccat gttatacctg
269 aattttacga tgagaaaaaa cttttttgata tcattggaag ccatgcacaa ttgcctacac
329 tttcaactgc gccaggttta gccgctaaaa tctacgtatc gcttattaat gaaggtgaaa
389 tggcatttag ttctgctgtt gaaggctcaa ccttcgatgc cacacttgtt caaagcactg
449 aggctgaacc aggcgttgtt gagattctcc aggtttcaca gccaattgtt aaggttggtg
509 cttcagctcc ggcaacacca gcaactccag caacaccagc aacaccatca aaaccagcaa
569 cacctgcagc tgtgtcaaca tccagtgccg gtgatgttgc aacaccagca gcatcacctt
629 ctgtggtgat tgccgagtca cctaacagtg ttgctgagtc acctgagagt tttggtgatg
689 cacctgctcc tgctccragc gcgtcttctc gtgccacatt cggattcatt ggtgctgtta
749 ttgcctttgc ttccattttt gtctctttgt aagcgttata taatgtattt ctagagagaa
809 aatgaaacaa tattgattgg aagtgataaa ttaattaatt tgttgctgct gggaaaaaat
869 caatgtaaca ccgttgattt gtgcttatgg gattcatttg ttcagattga
```

Figure 5

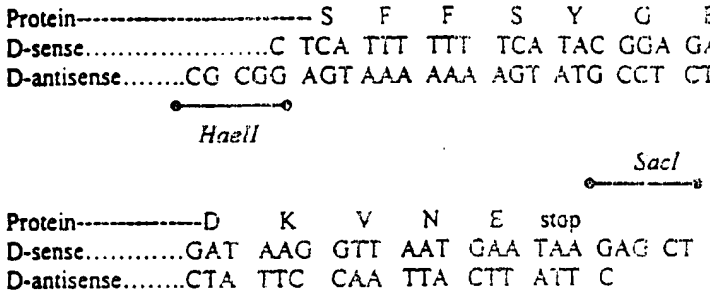
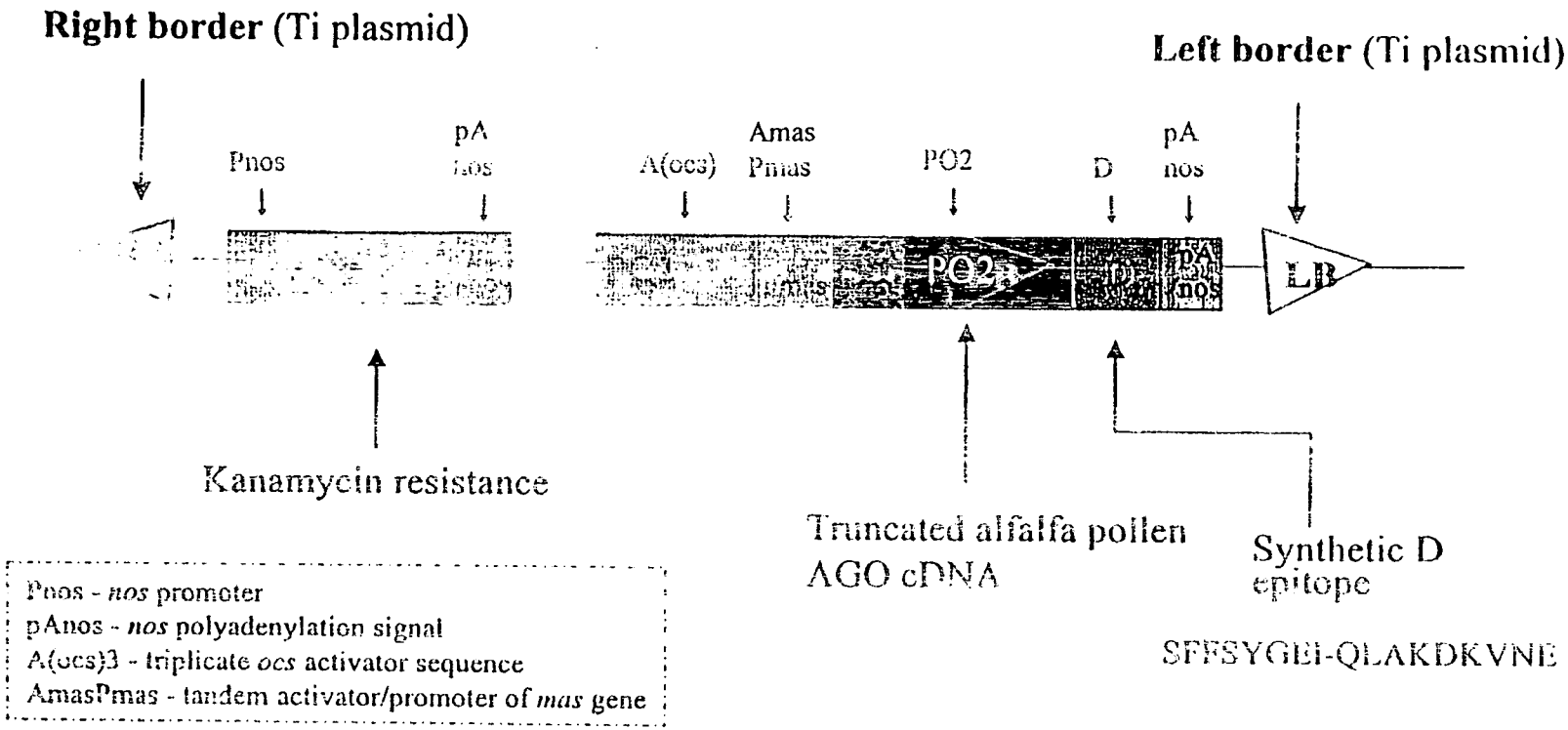

Protein---------- S F F S Y G E I Q L A K
D-sense.....C TCA TTT TTT TCA TAC GGA GAA ATT CAA TTG GCT AAG
D-antisense.....CG CGG AGT AAA AAA AGT ATG CCT CTT TAA GTT AAC CGA TTC
HaeII
SacI
Protein----------D K V N E stop
D-sense.....GAT AAG GTT AAT GAA TAA GAG CT
D-antisense.....CTA TTC CAA TTA CTT ATT C

Figure 6

NH2 -
- COOH
Truncated antarctic pollen AGP protein
• terminated at aa residue 237
• lacks membrane spanning domain
Synthetic D-epitope
• core SFFSYGEI sequence
• additional QLAKDKVNE sequence found to enhance antibody binding
• nucleotide sequence optimized for translation in alfalfa

Figure 7



Constitutive and universal expression of transgene under control of 'superpromoter'

Figure 8

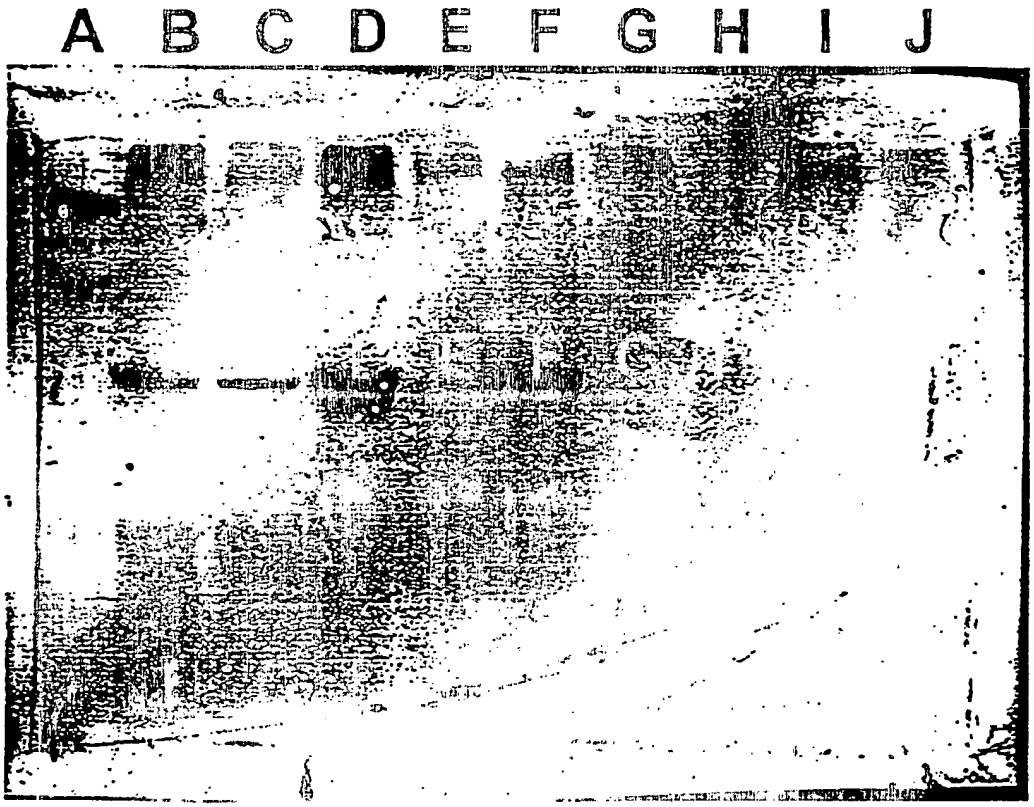


FIGURE 9

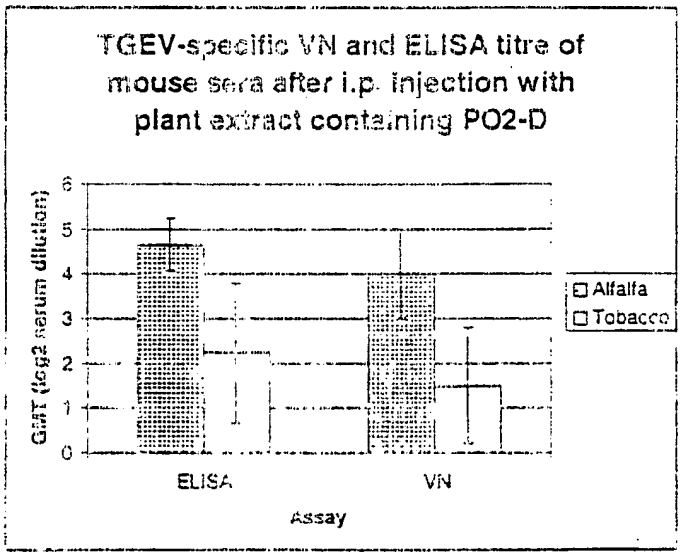

TGEV-specific VN and ELISA titre of mouse sera after i.p. injection with plant extract containing PO2-D
GMT (log2 serum dilution)
0
1
2
3
4
5
6
ELISA
VN
Assay
Alfalfa
Tobacco

Figure 10

SEQUENCE LISTING

<110> Bailey, Andrea
Erikson, Larry

<120> Method of Expressing Proteins and Peptides in Plants

<130> 08-884443US1

<140>
<141>

<150> US 60/151,147
<151> 1999-08-27

<160> 9

<170> PatentIn Ver. 2.1

<210> 1
<211> 938
<212> DNA
<213> Alfalfa pollen arabinogalactan

<400> 1
gaaagcaaga aaaaagaaca atgggtttga aaaacacctc tgttctgtgc ttggccatta 60
tgttgtcttt ctcttcggcg attaattgtt tggacataac caaaattatg ggacaatacc 120
cggaactctc cacgttcagc aaatacctaa ccgaaaccaa gttggctgac caaataaaca 180
gtggtaaagc cgttaccatt cttgcccttg acaacaaagc tactgcttct ctttctggta 240
aaccacttga cgccattaaa gccgttattg gaacccatgt tatacctgaa ttttacgatg 300
agaaaaaact ttttgatatc attggaagcc atgcacaatt gcctacactt tcaactgcgc 360
caggtttagc cgctaaaatc tacgtatcgc ttattaatga aggtgaaatg gcatttagtt 420
ctgctgttga aggctcaacc ttcgatgcca cacttgttca aagcactgag gctgaaccag 480
gcgttgttga gatttctccag gtttcacagc caattgttaa ggttggtgct tcagctccgg 540
caacaccagc aactccagca acaccagcaa caccatcaaa accagcaaca cctgcagctg 600
tgtcaacatc cagtgccggt gatgttgcaa caccagcagc atcaccttct gtggtgattg 660
ccgagtcacc taacagtgtt gctgagtcac ctgagagttt tggtgatgca cctgctcctg 720
ctcctagcgc ctcttctcgt gccacattcg gattcattgg tgctgttatt gcctttgctt 780
ccatttttgt ctctttgtaa gcgttatata atgtatttct agagagaaaa tgaaacaata 840
ttgattggaa gtgataaatt aattaatttg ttgctgctgg gttaaaatca atgtaacacc 900
gttgatttgt gcttatggga ttcatttgtt cagattga 938

<210> 2
<211> 259
<212> PRT
<213> Alfalfa pollen arabinogalactan

<400> 2
Met Gly Leu Lys Asn Thr Ser Val Leu Cys Leu Ala Ile Met Leu Ser
1 5 10 15

Phe Ser Ser Ala Ile Asn Cys Leu Asp Ile Gln Lys Ile Met Gly Gln
20 25 30

Tyr Pro Glu Leu Ser Thr Phe Ser Lys Tyr Leu Thr Glu Thr Lys Leu
35 40 45

Ala Asp Gln Ile Asn Ser Gly Lys Ala Val Thr Ile Leu Ala Leu Asp
50 55 60

Asn Lys Ala Ile Ala Ser Leu Ser Gly Lys Pro Leu Asp Ala Ile Lys
65 70 75 80

Ala Val Ile Gly Thr His Val Ile Pro Glu Phe Tyr Asp Glu Lys Lys
85 90 95

Leu Phe Asp Ile Ile Gly Ser His Ala Gln Leu Pro Thr Leu Ser Thr
100 105 110

Ala Pro Gly Leu Ala Ala Lys Ile Tyr Val Ser Leu Ile Asn Glu Gly
115 120 125

Glu Met Ala Phe Ser Ser Ala Val Glu Gly Ser Thr Phe Asp Ala Thr
130 135 140

Leu Val Gln Ser Thr Glu Ala Glu Pro Gly Val Val Glu Ile Leu Gln
145 150 155 160

Val Ser Gln Pro Ile Val Lys Val Gly Ala Ser Ala Pro Ala Thr Pro
165 170 175

Ala Thr Pro Ala Thr Pro Ala Thr Pro Ser Lys Pro Ala Thr Pro Ala
180 185 190

Ala Val Ser Thr Ser Ser Ala Gly Asp Val Ala Thr Pro Ala Ala Ser
195 200 205

Pro Ser Val Val Ile Ala Glu Ser Pro Asn Ser Val Ala Glu Ser Pro
210 215 220

Glu Ser Phe Gly Asp Ala Pro Ala Pro Ala Pro Ser Ala Ser Ser Arg
225 230 235 240

Ala Thr Phe Gly Phe Ile Gly Ala Val Ile Ala Phe Ala Ser Ile Phe
245 250 255

Val Ser Leu

<210> 3
<211> 254
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: AGP-D-epitope
of TGEV fusion protein

<400> 3
Met Gly Leu Lys Asn Thr Ser Val Leu Cys Leu Ala Ile Met Leu Ser
1 5 10 15

Phe Ser Ser Ala Ile Asn Cys Leu Asp Ile Thr Lys Ile Met Gly Gln
20 25 30

Tyr Pro Glu Leu Ser Thr Phe Ser Lys Tyr Leu Thr Glu Thr Lys Leu
35 40 45

Ala Asp Gln Ile Asn Ser Gly Lys Ala Val Thr Ile Leu Ala Leu Asp
50 55 60

Asn Lys Ala Ile Ala Ser Leu Ser Gly Lys Pro Leu Asp Ala Ile Lys
65 70 75 80

Ala Val Ile Gly Thr His Val Ile Pro Glu Phe Tyr Asp Glu Lys Lys
85 90 95

Leu Phe Asp Ile Ile Gly Ser His Ala Gln Leu Pro Thr Leu Ser Thr
100 105 110

Ala Pro Gly Leu Ala Ala Lys Ile Tyr Val Ser Leu Ile Asn Glu Gly
115 120 125

Glu Met Ala Phe Ser Ser Ala Val Glu Gly Ser Thr Phe Asp Ala Thr
130 135 140

Leu Val Gln Ser Thr Glu Ala Glu Pro Gly Val Val Glu Ile Leu Gln
145 150 155 160

Val Ser Gln Pro Ile Val Lys Val Gly Ala Ser Ala Pro Ala Thr Pro
165 170 175

Ala Thr Pro Ala Thr Pro Ala Thr Pro Ser Lys Pro Ala Thr Pro Ala
180 185 190

Ala Val Ser Thr Ser Ser Ala Gly Asp Val Ala Thr Pro Ala Ala Ser
195 200 205

Pro Ser Val Val Ile Ala Glu Ser Pro Asn Ser Val Ala Glu Ser Pro
210 215 220

Glu Ser Phe Gly Asp Ala Pro Ala Pro Ala Pro Ser Ala Ser Phe Phe
225 230 235 240

Ser Tyr Gly Glu Ile Gln Leu Ala Lys Asp Lys Val Asn Glu
245 250

<210> 4
<211> 950
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:truncated alfalfa pollen AGP

<400> 4
caatactacg tggaaagcaa gaaaaaagaa caatgggttt gaaaaacacc tctgttctgt 60
gcttggccat tatgttgtct ttctcttcgg cgattaattg tttggacata accaaaatta 120
tgggacaata cccggaactc tccacgttca gcaaatacct aaccgaaacc aagttggctg 180
accaaataaa cagtggtaaa gccgttacca ttcttgccct tgacaacaaa gctattgctt 240
ctcttttctgg taaaccactt gacgccatta aagccgttat tggaacccat gttatacctg 300
aattttacga tgagaaaaaa ctttttgata tcattggaag ccatgcacaa ttgcctacac 360
tttcaactgc gccaggttta gccgctaaaa tctacgtatc gcttattaat gaaggtgaaa 420
tggcatttag ttctgctgtt gaaggctcaa ccttcgatgc cacacttgtt caaagcactg 480
aggctgaacc aggcgttgtt gagattctcc aggtttcaca gccaattgtt aaggttggtg 540
cttcagctcc ggcaacacca gcaactccag cagcagcagc agccccatca aaaccagcaa 600
cacctgcagc tgtgccaaca tccagtgccg gtgatgttgc aacaccagca gcatcacctt 660
ctgtggtgat tgccgagtca cctaacagtg ttgctgagtc acctgagagt tttggtgatg 720
cacctgctcc tgctcctagc gccacttctc gtgccacatt cggattcatt ggtgctgtta 780
ttgcctttgc ttccatttct gtctcttctt aagcgctata taatgtattt ctagagagaa 840

aatgaaacaa tattgattgg aagtgataaa ttaattaatt tgttgctgct gggaaaaaat 900
caatgtaaca ccgttgattt gtgcttatgg gattcatttg tccagattga 950

<210> 5
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:sense oligo encoding modified TGEV epitope

<400> 5
ctcatttttt tcatacggag aaattcaatt ggccaaggat aaggttaatg aataagagct 60

<210> 6
<211> 60
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:antisense oligo encoding modified TGEV epitope

<400> 6
cgcggagtaa aaaaagtatg cctcttttaag ttaacgatt cctattccaa ttacttattc 60

<210> 7
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer PC2

<400> 7
cgctggccat atgggtttga aaaacacc 28

<210> 8
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:oligo for Bluescript

<400> 8
ggaaacagct atgaccatg 19

<210> 9
<211> 18
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:pBISNI primer

<400> 9
aagaacggaa tgcgcgtg 18

INTERNATIONAL SEARCH REPORT

Inter onal Application No
PCT/CA 00/00977

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/82 C12N15/63

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

BIOSIS, MEDLINE, EMBASE, SCISEARCH, BIOTECHNOLOGY ABS, EPO-Internal, WPI Data, STRAND

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE BIOSIS 'Online! BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; 1990 JENTOFT N: "WHY ARE PROTEINS O-GLYCOSYLATED" Database accession no. PREV199039081500 XP002155944 cited in the application abstract & TRENDS IN BIOCHEMICAL SCIENCES, vol. 15, no. 9, 1990, pages 291-294. ISSN: 0376-5067 -/---	

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

° Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
19 December 2000	22/01/2001

Name and mailing address of the ISA
European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer
Chakravarty, A

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/CA 00/00977

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	DATABASE BIOSIS 'Online! BIOSCIENCES INFORMATION SERVICE, PHILADELPHIA, PA, US; 1997 QIU XIAO ET AL.: "A new arabinogalactan protein-like gene expressed in the pollen of alfalfa." Database accession no. PREV199799565753 XP002155945 cited in the application abstract & PLANT SCIENCE (SHANNON), vol. 124, no. 1, 1997, pages 41-47, ISSN: 0168-9452	

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)